

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

- (A) NAME: National Starch and Chemical Investment Holding Corporation
- (B) STREET: Suite 27, 501 Silverside Road
- (C) CITY: Wilmington
- (D) STATE: Delaware
- (E) COUNTRY: USA
- (F) POSTAL CODE (ZIP): 19809

(ii) TITLE OF INVENTION: Improvements in or Relating to Starch Content of Plants

(iii) NUMBER OF SEQUENCES: 31

## (iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EP0)

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATGGACAAGG ATATGTATGA

20

## (2) INFORMATION FOR SEQ ID NO: 2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GGTTTCATGA CTTCTGAGCA

20

(2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

TGCTCAGAAG TCATGAAACC

20

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

TCCAGTCTCA ATATACGTCG

20

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

AGGAGTAGAT GGTCTGTCGA

20

## (2) INFORMATION FOR SEQ ID NO: 6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

TCATACATAT CCTTGTCCAT

20

## (2) INFORMATION FOR SEQ ID NO: 7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GGGTGACTTC AATGATGTAC

20

## (2) INFORMATION FOR SEQ ID NO: 8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GGTGTACATC ATTGAAGTCA

20

## (2) INFORMATION FOR SEQ ID NO: 9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

AATTACTGGC TCCGTACTAC

20

(2) INFORMATION FOR SEQ ID NO: 10:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CATTCCAACG TGC GACTCAT

20

(2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

TACCGGTAAT CTAGGTGTTG

20

(2) INFORMATION FOR SEQ ID NO: 12:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GGACCTTGGT TTAGATCCAA

20

## (2) INFORMATION FOR SEQ ID NO: 13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

ATGAGTCGCA CGTTGGAATG

20

## (2) INFORMATION FOR SEQ ID NO: 14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

CAACACCTAG ATTACCGGTA

20

## (2) INFORMATION FOR SEQ ID NO: 15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

TTAGTTGCGT CAGTTCTCAC

20

## (2) INFORMATION FOR SEQ ID NO: 16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

AATATCTATC TCAGCCGGAG

20

(2) INFORMATION FOR SEQ ID NO: 17:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

ATCTTAGATA GTCTGCATCA

20

(2) INFORMATION FOR SEQ ID NO: 18:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

TGGTTGTTCC CTGGAATTAC

20

(2) INFORMATION FOR SEQ ID NO: 19:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

TGCAAGGACC GTGACATCAA

20

## (2) INFORMATION FOR SEQ ID NO: 20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

CTTTATCTAT TAAAGACTTC

20

## (2) INFORMATION FOR SEQ ID NO: 21:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

CAAAAAAGTT TGTGACATGG

20

## (2) INFORMATION FOR SEQ ID NO: 22:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

TCACTTTTTC CAATGCTAAT

20

## (2) INFORMATION FOR SEQ ID NO: 23:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

TCTCATGCAA TGGAACCGAC

20

(2) INFORMATION FOR SEQ ID NO: 24:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

CAGATGTCCT GACTCGGAAT

20

(2) INFORMATION FOR SEQ ID NO: 25:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

ATTCCGAGTC AGGACATCTG

20

(2) INFORMATION FOR SEQ ID NO: 26:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:



CGCATTCTC GCTATTGCTT

20

## (2) INFORMATION FOR SEQ ID NO: 27:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

CACAGGCCCA AGTGAAGAAT

20

## (2) INFORMATION FOR SEQ ID NO: 28:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2588 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION:21..2531

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

CTCTCTAACT TCTCAGCGAA ATG GGA CAC TAC ACC ATA TCA GGA ATA CGT 50  
 Met Gly His Tyr Thr Ile Ser Gly Ile Arg  
 1 5 10

TTT CCT TGT GCT CCA CTC TGC AAA TCT CAA TCT ACC GGC TTC CAT GGC 98  
 Phe Pro Cys Ala Pro Leu Cys Lys Ser Gln Ser Thr Gly Phe His Gly  
 15 20 25

TAT CGG AGG ACC TCC TCT TGC CTT TCC TTC AAC TTC AAG GAG GCG TTT 146  
 Tyr Arg Arg Thr Ser Ser Cys Leu Ser Phe Asn Phe Lys Glu Ala Phe  
 30 35 40

TCT AGG AGG GTC TTC TCT GGA AAG TCA TCT CAT GAA TCT GAC TCC TCA 194  
 Ser Arg Arg Val Phe Ser Gly Lys Ser Ser His Glu Ser Asp Ser Ser  
 45 50 55

AAT GTA ATG GTC ACT GCT TCT AAA AGA GTC CTT CCT GAT GGT CGG ATT	242
Asn Val Met Val Thr Ala Ser Lys Arg Val Leu Pro Asp Gly Arg Ile	
60 65 70	
GAA TGC TAT TCT TCT TCA ACA GAT CAA TTG GAA GCC CCT GGC ACA GTT	290
Glu Cys Tyr Ser Ser Ser Thr Asp Gln Leu Glu Ala Pro Gly Thr Val	
75 80 85 90	
TCA GAA GAA TCC CAG GTG CTT ACT GAT GTT GAG AGT CTC ATT ATG GAT	338
Ser Glu Glu Ser Gln Val Leu Thr Asp Val Glu Ser Leu Ile Met Asp	
95 100 105	
GAT AAG ATT GTT GAA GAT GAA GTA AAT AAA GAA TCT GTT CCA ATG CGG	386
Asp Lys Ile Val Glu Asp Glu Val Asn Lys Glu Ser Val Pro Met Arg	
110 115 120	
GAG ACA GTT AGC ATC AGA AAA ATT GGA TCT AAA CCA AGG TCC ATT CCT	434
Glu Thr Val Ser Ile Arg Lys Ile Gly Ser Lys Pro Arg Ser Ile Pro	
125 130 135	
CCA CCC GGC AGA GGG CAA AGA ATA TAT GAC ATA GAT CCA AGC TTG ACA	482
Pro Pro Gly Arg Gly Gln Arg Ile Tyr Asp Ile Asp Pro Ser Leu Thr	
140 145 150	
GGC TTT CGT CAA CAC CTA GAT TAC CGG TAT TCA CAG TAC AAA AGA CTC	530
Gly Phe Arg Gln His Leu Asp Tyr Arg Tyr Ser Gln Tyr Lys Arg Leu	
155 160 165 170	
CGA GAA GAA ATT GAC AAG TAT GAA GGT AGT CTG GAT GCA TTT TCT CGT	578
Arg Glu Glu Ile Asp Lys Tyr Glu Gly Ser Leu Asp Ala Phe Ser Arg	
175 180 185	
GGC TAT GAA AAG TTT GGT TTC TCA CGC AGT GAA ACA GGA ATA ACT TAT	626
Gly Tyr Glu Lys Phe Gly Phe Ser Arg Ser Glu Thr Gly Ile Thr Tyr	
190 195 200	
AGA GAG TGG GCA CCA GGA GCT ACG TGG GCT GCA TTG ATT GGA GAT TTC	674
Arg Glu Trp Ala Pro Gly Ala Thr Trp Ala Ala Leu Ile Gly Asp Phe	
205 210 215	
AAT AAC TGG AAT CCT AAT GCA GAT GTC ATG ACT CAG AAT GAG TGT GGT	722
Asn Asn Trp Asn Pro Asn Ala Asp Val Met Thr Gln Asn Glu Cys Gly	
220 225 230	
GTC TGG GAG ATC TTT TTG CCG AAT AAT GCA GAT GGT TCA CCA CCA ATT	770
Val Trp Glu Ile Phe Leu Pro Asn Asn Ala Asp Gly Ser Pro Pro Ile	

235	240	245	250	
CCC CAT GGT TCT CGA GTA AAG ATA CGC ATG GAT ACT CCA TCT GGC AAC				818
Pro His Gly Ser Arg Val Lys Ile Arg Met Asp Thr Pro Ser Gly Asn	255	260	265	
AAA GAT TCT ATT CCT GCT TGG ATC AAG TTC TCA GTT CAA GCA CCA GGT				866
Lys Asp Ser Ile Pro Ala Trp Ile Lys Phe Ser Val Gln Ala Pro Gly	270	275	280	
GAA CTC CCA TAT AAT GGC ATA TAC TAT GAT CCT CCC GAG GAG GAG AAG				914
Glu Leu Pro Tyr Asn Gly Ile Tyr Tyr Asp Pro Pro Glu Glu Glu Lys	285	290	295	
TAT GTG TTC AAA AAT CCT CAG CCA AAG AGA CCA AAA TCA CTT CGG ATT				962
Tyr Val Phe Lys Asn Pro Gln Pro Lys Arg Pro Lys Ser Leu Arg Ile	300	305	310	
TAT GAG TCG CAC GTT GGA ATG AGT AGT ACG GAG CCA GTA ATT AAC ACA				1010
Tyr Glu Ser His Val Gly Met Ser Ser Thr Glu Pro Val Ile Asn Thr	315	320	325	330
TAT GCC AAC TTT AGA GAT GAT GTG CTT CCT CGC ATC AAA AAG CTT GGC				1058
Tyr Ala Asn Phe Arg Asp Asp Val Leu Pro Arg Ile Lys Lys Leu Gly	335	340	345	
TAC AAT GCT GTT CAG CTC ATG GCT ATT CAA GAG CAT TCA TAT TAT GCT				1106
Tyr Asn Ala Val Gln Leu Met Ala Ile Gln Glu His Ser Tyr Tyr Ala	350	355	360	
AGT TTT GGG TAT CAC GTC ACA AAC TTT TAT GCA GCT AGC AGC CGA TTT				1154
Ser Phe Gly Tyr His Val Thr Asn Phe Tyr Ala Ala Ser Ser Arg Phe	365	370	375	
GGA ACT CCT GAT GAT TTA AAG TCT CTA ATA GAT AAA GCT CAC GAG TTA				1202
Gly Thr Pro Asp Asp Leu Lys Ser Leu Ile Asp Lys Ala His Glu Leu	380	385	390	
GGT CTT CTT GTT CTC ATG GAT ATT GTT CAT AGC CAT GCA TCA ACT AAT				1250
Gly Leu Leu Val Leu Met Asp Ile Val His Ser His Ala Ser Thr Asn	395	400	405	410
ACG TTG GAT GGG CTG AAT ATG TTT GAT GGT ACG GAT GGT CAC TAC TTT				1298
Thr Leu Asp Gly Leu Asn Met Phe Asp Gly Thr Asp Gly His Tyr Phe	415	420	425	

CAC TCT GGA CCA CGG GGT CAT CAT TGG ATG TGG GAC TCT CGC CTT TTC	1346
His Ser Gly Pro Arg Gly His His Trp Met Trp Asp Ser Arg Leu Phe	
430 435 440	
AAC TAT GGG AGC TGG GAG GTT CTA AGG TTT CTT CTT TCA AAT GCA AGG	1394
Asn Tyr Gly Ser Trp Glu Val Leu Arg Phe Leu Leu Ser Asn Ala Arg	
445 450 455	
TGG TGG TTG GAT GAG TAC AAG TTT GAT GGG TTC AGA TTT GAT GGG GTG	1442
Trp Trp Leu Asp Glu Tyr Lys Phe Asp Gly Phe Arg Phe Asp Gly Val	
460 465 470	
ACT TCA ATG ATG TAC ACC CAT CAT GGA TTG CAG GTA GAT TTT ACC GGC	1490
Thr Ser Met Met Tyr Thr His His Gly Leu Gln Val Asp Phe Thr Gly	
475 480 485 490	
AAC TAC AAT GAA TAC TTT GGA TAT GCA ACT GAT GTA GAT GCT GTG GTT	1538
Asn Tyr Asn Glu Tyr Phe Gly Tyr Ala Thr Asp Val Asp Ala Val Val	
495 500 505	
TAT TTG ATG CTG TTG AAT GAT ATG ATT CAT GGT CTC TTC CCA GAG GCT	1586
Tyr Leu Met Leu Leu Asn Asp Met Ile His Gly Leu Phe Pro Glu Ala	
510 515 520	
GTC ACC ATT GGT GAA GAT GTT AGT GGA ATG CCA ACA GTT TGC ATT CCG	1634
Val Thr Ile Gly Glu Asp Val Ser Gly Met Pro Thr Val Cys Ile Pro	
525 530 535	
GTT GAA GAT GGT GGT GTT GGC TTT GAT TAT CGT CTC CAC ATG GCT GTT	1682
Val Glu Asp Gly Gly Val Gly Phe Asp Tyr Arg Leu His Met Ala Val	
540 545 550	
GCT GAT AAA TGG GTT GAG ATT ATT CAG AAG AGA GAT GAA GAT TGG AAA	1730
Ala Asp Lys Trp Val Glu Ile Ile Gln Lys Arg Asp Glu Asp Trp Lys	
555 560 565 570	
ATG GGT GAC ATT GTA CAT ATG CTG ACC AAC AGG CGG TGG TTG GAA AAG	1778
Met Gly Asp Ile Val His Met Leu Thr Asn Arg Arg Trp Leu Glu Lys	
575 580 585	
TGT GTT TCT TAT GCT GAA AGT CAT GAC CAG GCC CTT GTT GGT GAC AAA	1826
Cys Val Ser Tyr Ala Glu Ser His Asp Gln Ala Leu Val Gly Asp Lys	
590 595 600	

ACT ATT GCA TTT TGG CTG ATG GAC AAG GAT ATG TAT GAC TTC ATG GCT Thr Ile Ala Phe Trp Leu Met Asp Lys Asp Met Tyr Asp Phe Met Ala 605 610 615	1874
CTT GAC AGA CCA TCT ACT CCT CTC ATA GAT CGT GGA GTA GCA TTG CAC Leu Asp Arg Pro Ser Thr Pro Leu Ile Asp Arg Gly Val Ala Leu His 620 625 630	1922
AAA ATG ATC AGG CTT ATT ACC ATG GGA TTA GGC GGA GAA GGA TAT TTG Lys Met Ile Arg Leu Ile Thr Met Gly Leu Gly Gly Glu Gly Tyr Leu 635 640 645 650	1970
AAT TTT ATG GGA AAT GAA TTT GGA CAC CCC GAG TGG ATT GAT TTT CCA Asn Phe Met Gly Asn Glu Phe Gly His Pro Glu Trp Ile Asp Phe Pro 655 660 665	2018
AGA GGT GAT CTA CAT CTT CCC AGT GGT AAA TTT GTT CCT GGG AAC AAT Arg Gly Asp Leu His Leu Pro Ser Gly Lys Phe Val Pro Gly Asn Asn 670 675 680	2066
TAC AGT TAT GAT AAA TGC CGG CGT AGG TTT GAT CTA GGC AAT TCA AAG Tyr Ser Tyr Asp Lys Cys Arg Arg Arg Phe Asp Leu Gly Asn Ser Lys 685 690 695	2114
CAT CTG AGA TAT CAT GGA ATG CAA GAG TTT GAT CAA GCA ATT CAG CAT His Leu Arg Tyr His Gly Met Gln Glu Phe Asp Gln Ala Ile Gln His 700 705 710	2162
CTT GAA GAA GCC TAT GGT TTC ATG ACT TCT GAG CAC CAA TAC ATA TCA Leu Glu Glu Ala Tyr Gly Phe Met Thr Ser Glu His Gln Tyr Ile Ser 715 720 725 730	2210
CGG AAG GAT GAA AGG GAT CGG ATC ATT GTC TTC GAG AGG GGA AAC CTC Arg Lys Asp Glu Arg Asp Arg Ile Ile Val Phe Glu Arg Gly Asn Leu 735 740 745	2258
GTT TTT GTA TTC AAT TTT CAT TGG ACT AGC AGC TAT TCG GAT TAC CGA Val Phe Val Phe Asn Phe His Trp Thr Ser Ser Tyr Ser Asp Tyr Arg 750 755 760	2306
GTT GGC TGC TTA AAG CCA GGA AAG TAC AAG ATA GTC TTG GAT TCA GAT Val Gly Cys Leu Lys Pro Gly Lys Tyr Lys Ile Val Leu Asp Ser Asp 765 770 775	2354
GAT CCT TTG TTT GGA GGC TTT GGC AGG CTT AGT CAT GAT GCA GAG CAC	2402

Asp Pro Leu Phe Gly Gly Phe Gly Arg Leu Ser His Asp Ala Glu His  
 780 785 790  
 TTC AGC TTT GAA GGG TGG TAC GAT AAC CGG CCT CGA TCC TTC ATG GTG 2450  
 Phe Ser Phe Glu Gly Trp Tyr Asp Asn Arg Pro Arg Ser Phe Met Val  
 795 800 805 810  
 TAC ACA CCA TGT AGA ACA GCA GTG GTC TAT GCT TTA GTG GAG GAT GAA 2498  
 Tyr Thr Pro Cys Arg Thr Ala Val Val Tyr Ala Leu Val Glu Asp Glu  
 815 820 825  
 GTG GAG AAT GAA TTG GAA CCT GTC GCC GGT TAA GATATATCTT AACAAACAGGT 2551  
 Val Glu Asn Glu Leu Glu Pro Val Ala Gly \*  
 830 835  
 TCTGAAGCAG GAATGCCATT ATTGATCTTC CTATGTT 2588

## (2) INFORMATION FOR SEQ ID NO: 29:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 837 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

Met Gly His Tyr Thr Ile Ser Gly Ile Arg Phe Pro Cys Ala Pro Leu  
 1 5 10 15  
 Cys Lys Ser Gln Ser Thr Gly Phe His Gly Tyr Arg Arg Thr Ser Ser  
 20 25 30  
 Cys Leu Ser Phe Asn Phe Lys Glu Ala Phe Ser Arg Arg Val Phe Ser  
 35 40 45  
 Gly Lys Ser Ser His Glu Ser Asp Ser Ser Asn Val Met Val Thr Ala  
 50 55 60  
 Ser Lys Arg Val Leu Pro Asp Gly Arg Ile Glu Cys Tyr Ser Ser Ser  
 65 70 75 80  
 Thr Asp Gln Leu Glu Ala Pro Gly Thr Val Ser Glu Glu Ser Gln Val  
 85 90 95

Leu Thr Asp Val Glu Ser Leu Ile Met Asp Asp Lys Ile Val Glu Asp  
 100 105 110

Glu Val Asn Lys Glu Ser Val Pro Met Arg Glu Thr Val Ser Ile Arg  
 115 120 125

Lys Ile Gly Ser Lys Pro Arg Ser Ile Pro Pro Pro Gly Arg Gly Gln  
 130 135 140

Arg Ile Tyr Asp Ile Asp Pro Ser Leu Thr Gly Phe Arg Gln His Leu  
 145 150 155 160

Asp Tyr Arg Tyr Ser Gln Tyr Lys Arg Leu Arg Glu Glu Ile Asp Lys  
 165 170 175

Tyr Glu Gly Ser Leu Asp Ala Phe Ser Arg Gly Tyr Glu Lys Phe Gly  
 180 185 190

Phe Ser Arg Ser Glu Thr Gly Ile Thr Tyr Arg Glu Trp Ala Pro Gly  
 195 200 205

Ala Thr Trp Ala Ala Leu Ile Gly Asp Phe Asn Asn Trp Asn Pro Asn  
 210 215 220

Ala Asp Val Met Thr Gln Asn Glu Cys Gly Val Trp Glu Ile Phe Leu  
 225 230 235 240

Pro Asn Asn Ala Asp Gly Ser Pro Pro Ile Pro His Gly Ser Arg Val  
 245 250 255

Lys Ile Arg Met Asp Thr Pro Ser Gly Asn Lys Asp Ser Ile Pro Ala  
 260 265 270

Trp Ile Lys Phe Ser Val Gln Ala Pro Gly Glu Leu Pro Tyr Asn Gly  
 275 280 285

Ile Tyr Tyr Asp Pro Pro Glu Glu Lys Tyr Val Phe Lys Asn Pro  
 290 295 300

Gln Pro Lys Arg Pro Lys Ser Leu Arg Ile Tyr Glu Ser His Val Gly  
 305 310 315 320

Met Ser Ser Thr Glu Pro Val Ile Asn Thr Tyr Ala Asn Phe Arg Asp  
 325 330 335

Asp Val Leu Pro Arg Ile Lys Lys Leu Gly Tyr Asn Ala Val Gln Leu  
 340 345 350

Met Ala Ile Gln Glu His Ser Tyr Tyr Ala Ser Phe Gly Tyr His Val  
 355 360 365

Thr Asn Phe Tyr Ala Ala Ser Ser Arg Phe Gly Thr Pro Asp Asp Leu  
 370 375 380

Lys Ser Leu Ile Asp Lys Ala His Glu Leu Gly Leu Leu Val Leu Met  
 385 390 395 400

Asp Ile Val His Ser His Ala Ser Thr Asn Thr Leu Asp Gly Leu Asn  
 405 410 415

Met Phe Asp Gly Thr Asp Gly His Tyr Phe His Ser Gly Pro Arg Gly  
 420 425 430

His His Trp Met Trp Asp Ser Arg Leu Phe Asn Tyr Gly Ser Trp Glu  
 435 440 445

Val Leu Arg Phe Leu Leu Ser Asn Ala Arg Trp Trp Leu Asp Glu Tyr  
 450 455 460

Lys Phe Asp Gly Phe Arg Phe Asp Gly Val Thr Ser Met Met Tyr Thr  
 465 470 475 480

His His Gly Leu Gln Val Asp Phe Thr Gly Asn Tyr Asn Glu Tyr Phe  
 485 490 495

Gly Tyr Ala Thr Asp Val Asp Ala Val Val Tyr Leu Met Leu Leu Asn  
 500 505 510

Asp Met Ile His Gly Leu Phe Pro Glu Ala Val Thr Ile Gly Glu Asp  
 515 520 525

Val Ser Gly Met Pro Thr Val Cys Ile Pro Val Glu Asp Gly Gly Val  
 530 535 540

Gly Phe Asp Tyr Arg Leu His Met Ala Val Ala Asp Lys Trp Val Glu  
 545 550 555 560

Ile Ile Gln Lys Arg Asp Glu Asp Trp Lys Met Gly Asp Ile Val His  
 565 570 575



Met Leu Thr Asn Arg Arg Trp Leu Glu Lys Cys Val Ser Tyr Ala Glu  
580 585 590

Ser His Asp Gln Ala Leu Val Gly Asp Lys Thr Ile Ala Phe Trp Leu  
595 600 605

Met Asp Lys Asp Met Tyr Asp Phe Met Ala Leu Asp Arg Pro Ser Thr  
610 615 620

Pro Leu Ile Asp Arg Gly Val Ala Leu His Lys Met Ile Arg Leu Ile  
625 630 635 640

Thr Met Gly Leu Gly Gly Glu Gly Tyr Leu Asn Phe Met Gly Asn Glu  
645 650 655

Phe Gly His Pro Glu Trp Ile Asp Phe Pro Arg Gly Asp Leu His Leu  
660 665 670

Pro Ser Gly Lys Phe Val Pro Gly Asn Asn Tyr Ser Tyr Asp Lys Cys  
675 680 685

Arg Arg Arg Phe Asp Leu Gly Asn Ser Lys His Leu Arg Tyr His Gly  
690 695 700

Met Gln Glu Phe Asp Gln Ala Ile Gln His Leu Glu Glu Ala Tyr Gly  
705 710 715 720

Phe Met Thr Ser Glu His Gln Tyr Ile Ser Arg Lys Asp Glu Arg Asp  
725 730 735

Arg Ile Ile Val Phe Glu Arg Gly Asn Leu Val Phe Val Phe Asn Phe  
740 745 750

His Trp Thr Ser Ser Tyr Ser Asp Tyr Arg Val Gly Cys Leu Lys Pro  
755 760 765

Gly Lys Tyr Lys Ile Val Leu Asp Ser Asp Asp Pro Leu Phe Gly Gly  
770 775 780

Phe Gly Arg Leu Ser His Asp Ala Glu His Phe Ser Phe Glu Gly Trp  
785 790 795 800

Tyr Asp Asn Arg Pro Arg Ser Phe Met Val Tyr Thr Pro Cys Arg Thr  
805 810 815

Ala Val Val Tyr Ala Leu Val Glu Asp Glu Val Glu Asn Glu Leu Glu

820

825

830

Pro Val Ala Gly \*  
835

## (2) INFORMATION FOR SEQ ID NO: 30:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2805 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 131..2677

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

AGTGAATTCG AGCTCGGTAC CCGGGGATCC GATTCGCATT TCTCGCTATT GCTTTCCGTT	60
TATTTCCATA TATAAAATAT CAAATCTAAT CACTTGCGCC ATTTCTATCT CTCTCCAAAC	120
TCTCACCGAA ATG GTA TAC TAC ACT GTA TCA GGC ATA CGT TTT CCT TGT	169
Met Val Tyr Tyr Thr Val Ser Gly Ile Arg Phe Pro Cys	
840 845 850	
GCA CCT TCA CTC TAC AAA TCT CAG CTC ACC AGC TTC CAT GGC GGT CGA	217
Ala Pro Ser Leu Tyr Lys Ser Gln Leu Thr Ser Phe His Gly Gly Arg	
855 860 865	
AGG ACC TCT TCT GGC CTT TCC TTC CTC TTG AAG AAG GAG CTG TTT CCT	265
Arg Thr Ser Ser Gly Leu Ser Phe Leu Leu Lys Lys Glu Leu Phe Pro	
870 875 880	
CGG AAG ATC TTT GCT GGA AAG TCC TCT TAT GAA TCT GAC TCC TCA AAT	313
Arg Lys Ile Phe Ala Gly Lys Ser Ser Tyr Glu Ser Asp Ser Ser Asn	
885 890 895	
TTA ACT GTC TCT GCA TCT GAG AAG GTC CTT GTT CCT GAT GAT CAG ATT	361
Leu Thr Val Ser Ala Ser Glu Lys Val Leu Val Pro Asp Asp Gln Ile	
900 905 910	
GAT GGC TCT TCT TCT TCA ACA TAT CAA TTA GAA ACC ACT GGC ACA GTT	409

Asp Gly Ser Ser Ser Ser Thr Tyr Gln Leu Glu Thr Thr Gly Thr Val	
915 920 925 930	
TTG GAG GAA TCC CAG GTT CTT GGT GAT GCA GAG AGT CTT GTG ATG GAA	457
Leu Glu Glu Ser Gln Val Leu Gly Asp Ala Glu Ser Leu Val Met Glu	
935 940 945	
GAT GAT AAG AAT GTT GAG GAG GAT GAA GTA AAA AAA GAG TCG GTT CCA	505
Asp Asp Lys Asn Val Glu Glu Asp Glu Val Lys Lys Glu Ser Val Pro	
950 955 960	
TTG CAT GAG ACA ATT AGC ATT GGA AAA AGT GAA TCT AAA CCA AGG TCC	553
Leu His Glu Thr Ile Ser Ile Gly Lys Ser Glu Ser Lys Pro Arg Ser	
965 970 975	
ATT CCT CCA CCT GGC AGT GGG CAG AGA ATA TAT GAC ATA GAT CCA AGC	601
Ile Pro Pro Pro Gly Ser Gly Gln Arg Ile Tyr Asp Ile Asp Pro Ser	
980 985 990	
TTG GCA GGT TTC CGT CAG CAT CTT GAC TAC CGA TAT TCA CAG TAC AAA	649
Leu Ala Gly Phe Arg Gln His Leu Asp Tyr Arg Tyr Ser Gln Tyr Lys	
995 1000 1005 1010	
AGG CTG CGT GAG GAA ATT GAC AAG TAT GAA GGT GGT TTG GAT GCA TTC	697
Arg Leu Arg Glu Glu Ile Asp Lys Tyr Glu Gly Gly Leu Asp Ala Phe	
1015 1020 1025	
TCT CGT GGA TTT GAA AAG TTT GGT TTC TTA CGC AGT GAA ACA GGA ATA	745
Ser Arg Gly Phe Glu Lys Phe Gly Phe Leu Arg Ser Glu Thr Gly Ile	
1030 1035 1040	
ACT TAT AGG GAA TGG GCA CCT GGA GCT ACG TGG GCT GCA CTT ATT GGA	793
Thr Tyr Arg Glu Trp Ala Pro Gly Ala Thr Trp Ala Ala Leu Ile Gly	
1045 1050 1055	
GAT TTC AAC AAT TGG AAT CCT AAT GCA GAT GTC ATG ACT CGG AAT GAG	841
Asp Phe Asn Asn Trp Asn Pro Asn Ala Asp Val Met Thr Arg Asn Glu	
1060 1065 1070	
TTT GGT GTC TGG GAG ATT TTT TTG CCA AAT AAC GCA GAT GGT TCA CCA	889
Phe Gly Val Trp Glu Ile Phe Leu Pro Asn Asn Ala Asp Gly Ser Pro	
1075 1080 1085 1090	
CCA ATT CCT CAT GGT TCT CGA GTA AAG ATA CGC ATG GAT ACT CCA TCT	937
Pro Ile Pro His Gly Ser Arg Val Lys Ile Arg Met Asp Thr Pro Ser	

1095	1100	1105	
GGC ATC AAA GAT TCA ATT CCT GCT TGG ATC AAG TTC TCA GTT CAG GCA Gly Ile Lys Asp Ser Ile Pro Ala Trp Ile Lys Phe Ser Val Gln Ala 1110 1115 1120			985
CCT GGT GAA ATC CCA TAC AAT GCC ATA TAC TAT GAT CCA CCA AAG GAG Pro Gly Glu Ile Pro Tyr Asn Ala Ile Tyr Tyr Asp Pro Pro Lys Glu 1125 1130 1135			1033
GAG AAG TAT GTG TTC AAA CAT CCT CAG CCA AAG AGA CCA AAA TCA CTT Glu Lys Tyr Val Phe Lys His Pro Gln Pro Lys Arg Pro Lys Ser Leu 1140 1145 1150			1081
AGG ATT TAT GAA TCT CAT GTT GGG ATG AGT AGT ATG GAG CCA ATA ATT Arg Ile Tyr Glu Ser His Val Gly Met Ser Ser Met Glu Pro Ile Ile 1155 1160 1165 1170			1129
AAC ACA TAT GCC AAC TTT AGA GAT GAT ATG CTT CCT CGC ATC AAA AAG Asn Thr Tyr Ala Asn Phe Arg Asp Asp Met Leu Pro Arg Ile Lys Lys 1175 1180 1185			1177
CTT GGC TAC AAT GCT GTT CAG ATC ATG GCT ATT CAA GAG CAT TCC TAT Leu Gly Tyr Asn Ala Val Gln Ile Met Ala Ile Gln Glu His Ser Tyr 1190 1195 1200			1225
TAT GCT AGT TTT GGG TAC CAT GTC ACA AAC TTT TTT GCA CCT AGC AGC Tyr Ala Ser Phe Gly Tyr His Val Thr Asn Phe Phe Ala Pro Ser Ser 1205 1210 1215			1273
CGA TTT GGA ACT CCT GAT GAT TTG AAG TCT TTA ATA GAT AAA GCT CAT Arg Phe Gly Thr Pro Asp Asp Leu Lys Ser Leu Ile Asp Lys Ala His 1220 1225 1230			1321
GAG TTA GGG CTG CTT GTT CTC ATG GAT ATT GTT CAT AGC CAT GCG TCA Glu Leu Gly Leu Leu Val Leu Met Asp Ile Val His Ser His Ala Ser 1235 1240 1245 1250			1369
AAT AAT ACG TTG GAT GGG CTG AAC ATG TTT GAT GGT ACG GAT AGT CAC Asn Asn Thr Leu Asp Gly Leu Asn Met Phe Asp Gly Thr Asp Ser His 1255 1260 1265			1417
TAC TTC CAC TCC GGA TCA CGG GGT CAT CAT TGG TTG TGG GAC TCT CGC Tyr Phe His Ser Gly Ser Arg Gly His His Trp Leu Trp Asp Ser Arg 1270 1275 1280			1465

CTT TTC AAC TAT GGA AGC TGG GAG GTG CTA AGA TTT CTT CTT TCA AAT Leu Phe Asn Tyr Gly Ser Trp Glu Val Leu Arg Phe Leu Leu Ser Asn 1285 1290 1295	1513
GCA AGA TGG TGG TTG GAA GAG TAC AGG TTT GAT GGT TTT AGA TTT GAT Ala Arg Trp Trp Leu Glu Glu Tyr Arg Phe Asp Gly Phe Arg Phe Asp 1300 1305 1310	1561
GGG GTG ACT TCC ATG ATG TAC ACT CCC CAT GGG TTG CAG GTA GCT TTT Gly Val Thr Ser Met Met Tyr Thr Pro His Gly Leu Gln Val Ala Phe 1315 1320 1325 1330	1609
ACT GGC AAC TAC AAT GAG TAC TTT GGA TAT GCA ACT GAT GTA GAT GCT Thr Gly Asn Tyr Asn Glu Tyr Phe Gly Tyr Ala Thr Asp Val Asp Ala 1335 1340 1345	1657
GTG ATT TAT TTG ATG CTT GTG AAT GAT ATG ATT CAC GGT CTT TTC CCT Val Ile Tyr Leu Met Leu Val Asn Asp Met Ile His Gly Leu Phe Pro 1350 1355 1360	1705
GAG GCT GTT ACC ATT GGT GAA GAT GTT AGC GGA AAG CCA ACA TTT TGC Glu Ala Val Thr Ile Gly Glu Asp Val Ser Gly Lys Pro Thr Phe Cys 1365 1370 1375	1753
ATT CCA GTG GAA GAT GGT GGT GTT GGA TTT GAT TAC CGT CTC CAC ATG Ile Pro Val Glu Asp Gly Gly Val Gly Phe Asp Tyr Arg Leu His Met 1380 1385 1390	1801
GCC ATT GCC GAT AAA TGG ATT GAG ATT CTT AAG AAG AGA GAT GAG GAC Ala Ile Ala Asp Lys Trp Ile Glu Ile Leu Lys Lys Arg Asp Glu Asp 1395 1400 1405 1410	1849
TGG AAA ATG GGT GAC ATT GTG CAT ACA CTC ACC AAC AGA AGG TGG TTG Trp Lys Met Gly Asp Ile Val His Thr Leu Thr Asn Arg Arg Trp Leu 1415 1420 1425	1897
GAA AAA TGT GTT GCT TAT GCT GAA AGT CAT GAC CAA GCT CTT GTT GGT Glu Lys Cys Val Ala Tyr Ala Glu Ser His Asp Gln Ala Leu Val Gly 1430 1435 1440	1945
GAC AAA ACT ATT GCA TTT TGG CTG ATG GAC AAG GAC ATG TAC GAC TTC Asp Lys Thr Ile Ala Phe Trp Leu Met Asp Lys Asp Met Tyr Asp Phe 1445 1450 1455	1993
ATG GCT CGT GAC AGA CCA TCT ACT CCT CTT ATA GAT CGT GGA ATA GCA	2041

Met	Ala	Arg	Asp	Arg	Pro	Ser	Thr	Pro	Leu	Ile	Asp	Arg	Gly	Ile	Ala	
1460						1465					1470					
TTG	CAC	AAA	ATG	ATC	AGG	CTT	ATT	ACC	ATG	GGC	TTA	GGC	GGA	GAA	GGA	2089
Leu	His	Lys	Met	Ile	Arg	Leu	Ile	Thr	Met	Gly	Leu	Gly	Gly	Glu	Gly	
1475					1480					1485					1490	
TAT	TTG	AAT	TTT	ATG	GGA	AAT	GAA	TTT	GGA	CAT	CCT	GAG	TGG	ATT	GAT	2137
Tyr	Leu	Asn	Phe	Met	Gly	Asn	Glu	Phe	Gly	His	Pro	Glu	Trp	Ile	Asp	
				1495					1500					1505		
TTT	CCA	AGA	GGG	GAT	CGA	CAT	CTG	CCC	AAT	GGT	AAA	GTA	ATT	CCA	GGG	2185
Phe	Pro	Arg	Gly	Asp	Arg	His	Leu	Pro	Asn	Gly	Lys	Val	Ile	Pro	Gly	
			1510					1515					1520			
AAC	AAC	CAC	AGT	TAT	GAT	AAA	TGC	CGT	CGT	AGA	TTT	GAT	CTA	GGT	GAT	2233
Asn	Asn	His	Ser	Tyr	Asp	Lys	Cys	Arg	Arg	Arg	Phe	Asp	Leu	Gly	Asp	
		1525				1530						1535				
GCA	GAC	TAT	CTA	AGA	TAT	CAT	GGA	ATG	CAA	GAG	TTT	GAT	CAG	GCA	ATG	2281
Ala	Asp	Tyr	Leu	Arg	Tyr	His	Gly	Met	Gln	Glu	Phe	Asp	Gln	Ala	Met	
	1540				1545					1550						
CAA	CAT	CTT	GAA	GAA	GCC	TAT	GGT	TTC	ATG	ACT	TCT	GAG	CAC	CAG	TAT	2329
Gln	His	Leu	Glu	Glu	Ala	Tyr	Gly	Phe	Met	Thr	Ser	Glu	His	Gln	Tyr	
1555					1560					1565					1570	
ATA	TCA	CGG	AAG	GAT	GAA	GGA	GAT	CGG	ATC	ATT	GTC	TTT	GAG	AGG	GGA	2377
Ile	Ser	Arg	Lys	Asp	Glu	Gly	Asp	Arg	Ile	Ile	Val	Phe	Glu	Arg	Gly	
			1575					1580					1585			
AAC	CTT	GTT	TTT	GTA	TTC	AAC	TTT	CAT	TGG	ACT	AAC	AGC	TAT	TCA	GAT	2425
Asn	Leu	Val	Phe	Val	Phe	Asn	Phe	His	Trp	Thr	Asn	Ser	Tyr	Ser	Asp	
		1590						1595				1600				
TAC	CGA	GTT	GGC	TGC	TTC	AAG	TCA	GGA	AAG	TAC	AAG	ATT	GTT	TTG	GAC	2473
Tyr	Arg	Val	Gly	Cys	Phe	Lys	Ser	Gly	Lys	Tyr	Lys	Ile	Val	Leu	Asp	
	1605					1610						1615				
TCG	GAT	GAT	GGC	TTG	TTT	GGA	GGC	TTC	AAC	AGG	CTT	AGT	CAT	GAT	GCC	2521
Ser	Asp	Asp	Gly	Leu	Phe	Gly	Gly	Phe	Asn	Arg	Leu	Ser	His	Asp	Ala	
	1620				1625					1630						
GAG	CAC	TTC	ACC	TTT	GAC	GGG	TGG	TAT	GAT	AAC	CGG	CCT	CGG	TCC	TTC	2569
Glu	His	Phe	Thr	Phe	Asp	Gly	Trp	Tyr	Asp	Asn	Arg	Pro	Arg	Ser	Phe	

1635	1640	1645	1650	
ATG GTA TAT GCA CCA TCT AGG ACA GCA GTG GTC TAT GCT TTA GTA GAA				2617
Met Val Tyr Ala Pro Ser Arg Thr Ala Val Val Tyr Ala Leu Val Glu				
	1655	1660	1665	
GAT GAA GAG AAT GAA GCA GAG AAT GAA GTA GAA AGT GAA GTG AAA CCA				2665
Asp Glu Glu Asn Glu Ala Glu Asn Glu Val Glu Ser Glu Val Lys Pro				
	1670	1675	1680	
GCC TCC GGC TGA GATAGATATT TAGTAAGAGG ATCCCCTAAA GCAGGAATGG				2717
Ala Ser Gly *				
	1685			
TTAACCTGTG CATCTGCATT GAACGACGTA TATTGAGACT GGAAATCCAT ATGACTAGTA				2777
GATCCTCTAG AGTCGACCTG CAGGCATG				2805

## (2) INFORMATION FOR SEQ ID NO: 31:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 849 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

Met Val Tyr Tyr Thr Val Ser Gly Ile Arg Phe Pro Cys Ala Pro Ser			
1	5	10	15
Leu Tyr Lys Ser Gln Leu Thr Ser Phe His Gly Gly Arg Arg Thr Ser			
	20	25	30
Ser Gly Leu Ser Phe Leu Leu Lys Lys Glu Leu Phe Pro Arg Lys Ile			
	35	40	45
Phe Ala Gly Lys Ser Ser Tyr Glu Ser Asp Ser Ser Asn Leu Thr Val			
	50	55	60
Ser Ala Ser Glu Lys Val Leu Val Pro Asp Asp Gln Ile Asp Gly Ser			
	65	70	75
Ser Ser Ser Thr Tyr Gln Leu Glu Thr Thr Gly Thr Val Leu Glu Glu			

				85					90					95	
Ser	Gln	Val	Leu	Gly	Asp	Ala	Glu	Ser	Leu	Val	Met	Glu	Asp	Asp	Lys
			100					105					110		
Asn	Val	Glu	Glu	Asp	Glu	Val	Lys	Lys	Glu	Ser	Val	Pro	Leu	His	Glu
		115					120					125			
Thr	Ile	Ser	Ile	Gly	Lys	Ser	Glu	Ser	Lys	Pro	Arg	Ser	Ile	Pro	Pro
	130					135					140				
Pro	Gly	Ser	Gly	Gln	Arg	Ile	Tyr	Asp	Ile	Asp	Pro	Ser	Leu	Ala	Gly
145					150					155					160
Phe	Arg	Gln	His	Leu	Asp	Tyr	Arg	Tyr	Ser	Gln	Tyr	Lys	Arg	Leu	Arg
				165					170					175	
Glu	Glu	Ile	Asp	Lys	Tyr	Glu	Gly	Gly	Leu	Asp	Ala	Phe	Ser	Arg	Gly
			180					185					190		
Phe	Glu	Lys	Phe	Gly	Phe	Leu	Arg	Ser	Glu	Thr	Gly	Ile	Thr	Tyr	Arg
		195					200					205			
Glu	Trp	Ala	Pro	Gly	Ala	Thr	Trp	Ala	Ala	Leu	Ile	Gly	Asp	Phe	Asn
	210					215					220				
Asn	Trp	Asn	Pro	Asn	Ala	Asp	Val	Met	Thr	Arg	Asn	Glu	Phe	Gly	Val
225				230						235					240
Trp	Glu	Ile	Phe	Leu	Pro	Asn	Asn	Ala	Asp	Gly	Ser	Pro	Pro	Ile	Pro
				245					250					255	
His	Gly	Ser	Arg	Val	Lys	Ile	Arg	Met	Asp	Thr	Pro	Ser	Gly	Ile	Lys
			260					265					270		
Asp	Ser	Ile	Pro	Ala	Trp	Ile	Lys	Phe	Ser	Val	Gln	Ala	Pro	Gly	Glu
		275					280					285			
Ile	Pro	Tyr	Asn	Ala	Ile	Tyr	Tyr	Asp	Pro	Pro	Lys	Glu	Glu	Lys	Tyr
	290					295					300				
Val	Phe	Lys	His	Pro	Gln	Pro	Lys	Arg	Pro	Lys	Ser	Leu	Arg	Ile	Tyr
305					310					315					320
Glu	Ser	His	Val	Gly	Met	Ser	Ser	Met	Glu	Pro	Ile	Ile	Asn	Thr	Tyr



			325				330				335				
Ala	Asn	Phe	Arg	Asp	Asp	Met	Leu	Pro	Arg	Ile	Lys	Lys	Leu	Gly	Tyr
			340					345				350			
Asn	Ala	Val	Gln	Ile	Met	Ala	Ile	Gln	Glu	His	Ser	Tyr	Tyr	Ala	Ser
			355					360				365			
Phe	Gly	Tyr	His	Val	Thr	Asn	Phe	Phe	Ala	Pro	Ser	Ser	Arg	Phe	Gly
			370					375				380			
Thr	Pro	Asp	Asp	Leu	Lys	Ser	Leu	Ile	Asp	Lys	Ala	His	Glu	Leu	Gly
			385					390				395			
Leu	Leu	Val	Leu	Met	Asp	Ile	Val	His	Ser	His	Ala	Ser	Asn	Asn	Thr
				405				410				415			
Leu	Asp	Gly	Leu	Asn	Met	Phe	Asp	Gly	Thr	Asp	Ser	His	Tyr	Phe	His
			420				425				430				
Ser	Gly	Ser	Arg	Gly	His	His	Trp	Leu	Trp	Asp	Ser	Arg	Leu	Phe	Asn
			435				440				445				
Tyr	Gly	Ser	Trp	Glu	Val	Leu	Arg	Phe	Leu	Leu	Ser	Asn	Ala	Arg	Trp
			450				455				460				
Trp	Leu	Glu	Glu	Tyr	Arg	Phe	Asp	Gly	Phe	Arg	Phe	Asp	Gly	Val	Thr
			465				470				475				
Ser	Met	Met	Tyr	Thr	Pro	His	Gly	Leu	Gln	Val	Ala	Phe	Thr	Gly	Asn
			485				490				495				
Tyr	Asn	Glu	Tyr	Phe	Gly	Tyr	Ala	Thr	Asp	Val	Asp	Ala	Val	Ile	Tyr
			500				505				510				
Leu	Met	Leu	Val	Asn	Asp	Met	Ile	His	Gly	Leu	Phe	Pro	Glu	Ala	Val
			515				520				525				
Thr	Ile	Gly	Glu	Asp	Val	Ser	Gly	Lys	Pro	Thr	Phe	Cys	Ile	Pro	Val
			530				535				540				
Glu	Asp	Gly	Gly	Val	Gly	Phe	Asp	Tyr	Arg	Leu	His	Met	Ala	Ile	Ala
			545				550				555				
Asp	Lys	Trp	Ile	Glu	Ile	Leu	Lys	Lys	Arg	Asp	Glu	Asp	Trp	Lys	Met

				565				570				575			
Gly	Asp	Ile	Val	His	Thr	Leu	Thr	Asn	Arg	Arg	Trp	Leu	Glu	Lys	Cys
580								585				590			
Val	Ala	Tyr	Ala	Glu	Ser	His	Asp	Gln	Ala	Leu	Val	Gly	Asp	Lys	Thr
595								600				605			
Ile	Ala	Phe	Trp	Leu	Met	Asp	Lys	Asp	Met	Tyr	Asp	Phe	Met	Ala	Arg
610								615				620			
Asp	Arg	Pro	Ser	Thr	Pro	Leu	Ile	Asp	Arg	Gly	Ile	Ala	Leu	His	Lys
625								630				635			
Met	Ile	Arg	Leu	Ile	Thr	Met	Gly	Leu	Gly	Gly	Glu	Gly	Tyr	Leu	Asn
				645				650				655			
Phe	Met	Gly	Asn	Glu	Phe	Gly	His	Pro	Glu	Trp	Ile	Asp	Phe	Pro	Arg
				660				665				670			
Gly	Asp	Arg	His	Leu	Pro	Asn	Gly	Lys	Val	Ile	Pro	Gly	Asn	Asn	His
675								680				685			
Ser	Tyr	Asp	Lys	Cys	Arg	Arg	Arg	Phe	Asp	Leu	Gly	Asp	Ala	Asp	Tyr
690								695				700			
Leu	Arg	Tyr	His	Gly	Met	Gln	Glu	Phe	Asp	Gln	Ala	Met	Gln	His	Leu
705								710				715			
Glu	Glu	Ala	Tyr	Gly	Phe	Met	Thr	Ser	Glu	His	Gln	Tyr	Ile	Ser	Arg
				725				730				735			
Lys	Asp	Glu	Gly	Asp	Arg	Ile	Ile	Val	Phe	Glu	Arg	Gly	Asn	Leu	Val
				740				745				750			
Phe	Val	Phe	Asn	Phe	His	Trp	Thr	Asn	Ser	Tyr	Ser	Asp	Tyr	Arg	Val
				755				760				765			
Gly	Cys	Phe	Lys	Ser	Gly	Lys	Tyr	Lys	Ile	Val	Leu	Asp	Ser	Asp	Asp
770								775				780			
Gly	Leu	Phe	Gly	Gly	Phe	Asn	Arg	Leu	Ser	His	Asp	Ala	Glu	His	Phe
785								790				795			
Thr	Phe	Asp	Gly	Trp	Tyr	Asp	Asn	Arg	Pro	Arg	Ser	Phe	Met	Val	Tyr

50

805

810

815

Ala Pro Ser Arg Thr Ala Val Val Tyr Ala Leu Val Glu Asp Glu Glu  
820 825 830

Asn Glu Ala Glu Asn Glu Val Glu Ser Glu Val Lys Pro Ala Ser Gly  
835 840 845 \*